Michela Eleuteri

ANALISI MATEMATICA

Appunti di algebra lineare

Sistemi lineari. Diagonalizzazione (teoria)

A Giulia

con la speranza che almeno nella matematica

non assomigli al papà ☺

Indice

CAPITOLO 1

Teoria dei sistemi lineari

1.1. Generalità

Sia $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$ una trasformazione lineare e sia $\mathbf{A}$ la matrice che la rappresenta. In questo capitolo andiamo a discutere l'esistenza di soluzioni per l'equazione $\mathcal{L}\mathbf{x} = \mathbf{b}$ o equivalentemente $\mathbf{A}\mathbf{x} = \mathbf{b}$.

❒ **Definizione 1.1.1.** *Per* SISTEMA LINEARE *si intende un insieme di equazioni algebriche di primo grado. Il numero delle equzioni o delle incognite può variare. Generalmente si scrivono a primo membro dell'equazione le incognite moltiplicate per i loro coefficienti, mentre a secondo membro si scrivono i termini noti; se questi sono tutti nulli il sistema si dice* OMOGENEO.
Per SOLUZIONE *di un sistema lineare si intende una n-upla di numeri che, sostituiti tutti ordinatamente alle incognite, soddisfano simultaneamente tutte le equazioni del sistema. Esistono sistemi* DETERMINATI, *cioè che ammettono una sola soluzione,* INDETERMINATI, *cioè che ammettono infinite soluzioni,* IMPOSSIBILI *cioè che non ammettono soluzioni.*

✎ **Esempio 1.1.2.**

$$\begin{cases} x + 3y + 5z = 0 \\ 4x - 3y + 7z = -1 \end{cases} \qquad \text{sistema di 2 equazioni in 3 incognite.}$$

✎ **Esempio 1.1.3.** *Il più semplice sistema lineare è costituito da un'equazione e un'incognita*

$$ax = b.$$

Sappiamo che se $a \neq 0$ allora esiste un'unica soluzione $x = \frac{b}{a}$; se $a = 0$ e $b = 0$ esistono infinite soluzioni mentre se $a = 0$ e $b \neq 0$ non esiste alcuna soluzione. Questo esempio sarà il prototipo che ci guiderà nella teoria dei sistemi lineari.

✎ **Esempio 1.1.4.** *Consideriamo i seguenti sistemi lineari*

$$\begin{cases} 3x + 2y = 4 \\ x - 2y = 2 \end{cases} \qquad \begin{cases} x - 2y = 4 \\ x - 2y = 2 \end{cases} \qquad \begin{cases} 3x + 4y = 1 \\ 6x + 8y = 2 \end{cases}$$

Il primo è determinato e dà come soluzione $x = \frac{3}{2}, y = -\frac{1}{4}$*, il secondo non ammtte soluzioni, il terzo ammette infinite soluzioni. Dal punto di vista geometrico (vedi Sezione 1.9.1, il primo rappresenta l'intersezione di due rette incidenti e quindi la soluzione è un punto, il secondo rappresenta due rette parallele e il terzo rappresenta la stessa retta.*

1.2. Sistemi di n equazioni in n incognite non omogenei

Consideriamo il caso di un sistema di n equazioni in n incognite non omogeneo, del tipo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{1.2.1}$$

La matrice $\mathbf{A}$ dei coefficienti è data da

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Introducendo i vettori colonna

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

delle incognite e dei termini noti rispettivamente, possiamo riscrivere il sistema 1.2.1 in forma matriciale come

$$\mathbf{Ax} = \mathbf{b}. \tag{1.2.2}$$

Per risolverlo, ricorriamo all'analogia con l'Esempio 1.1.3. Se det$\mathbf{A} \neq 0$, allora per il teorema sulla matrice inversa, esiste $\mathbf{A}^{-1}$: moltiplicando a sinistra per $\mathbf{A}^{-1}$ ambo i membri della (1.2.2) e utilizzando la proprietà associativa del prodotto di matrici, si ottiene

$$\mathbf{x} = (\mathbf{AA}^{-1})\mathbf{x} = \mathbf{A}^{-1}(\mathbf{Ax}) = \mathbf{A}^{-1}\mathbf{b}$$

e dunque

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Si osserva che, sviluppando i calcoli indicati e ricordando la formula per la matrice inversa si ha

$$x_i = (\mathbf{A}^{-1}b)_i = \frac{1}{\det\mathbf{A}}\sum_{j=1}^{n} A_{ji}b_j = \frac{1}{\det\mathbf{A}}\det\mathbf{B}_i$$

dove $\mathbf{B}_i$ è la matrice che si ottiene dalla matrice $\mathbf{A}$ sostituendo alla colonna i-esima la colonna dei termini noti, cioè

$$\mathbf{B}_i = \begin{pmatrix} a_{11} & a_{12} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_2 & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & b_n & \dots & a_{nn} \end{pmatrix}$$

Quindi

$$x_i = \frac{|\mathbf{B}_i|}{|\mathbf{A}|} \tag{1.2.3}$$

Abbiamo così dimostrato il seguente teorema.

Teorema 1.2.1. (teorema di Cramer) *Consideriamo il sistema di n equazioni in n incognite*

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

con $\mathbf{A}$ matrice di tipo (n, n), $\mathbf{b} \in \mathbb{R}^n$ assegnato e $\mathbf{x} \in \mathbb{R}^n$ incognito. Se $\det\mathbf{A} \neq 0$, allora il sistema è determinato ossia ha una sola soluzione data dalla formula (1.2.3).

✎ **Esempio 1.2.2.** *Sia dato il sistema lineare in due equazioni e due incognite*

$$\begin{cases} 3x + 2y = 4 \\ x - 2y = 2 \end{cases}$$

Allora

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & -2 \end{pmatrix}$$

dunque $\det\mathbf{A} = -8 \neq 0$ e si ha

$$x = \frac{\begin{vmatrix} 4 & 2 \\ 2 & -2 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 1 & -2 \end{vmatrix}} = \frac{3}{2} \qquad y = \frac{\begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 1 & -2 \end{vmatrix}} = -\frac{1}{4}.$$

✎ **Esempio 1.2.3.** *Discutere esistenza e unicità eventuale delle soluzioni del seguente sistema lineare non omogeneo in tre equazioni e tre incognite*

$$\begin{cases} 2x + y + z = 3 \\ x - y = 1 \\ y + z = -1 \end{cases}$$

e fornire un'interpretazione geometrica.

La matrice dei coefficienti è data da

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

mentre la colonna dei termini noti è data da

$$\mathbf{b} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$$

Si ha che det$\mathbf{A} = -2 \neq 0$ quindi dal Teorema di Cramer, il sistema ammette un'unica soluzione data da

$$x = \frac{\begin{vmatrix} 3 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{vmatrix}} = 2 \qquad y = x = \frac{\begin{vmatrix} 2 & 3 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{vmatrix}} = 1 \qquad z = x = \frac{\begin{vmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{vmatrix}} = -2$$

Dal punto di vista geometrico, si tratta dell'intersezione di 3 piani che si incontrano in un punto.

1.3. Sistemi omogenei di n equazioni in n incognite

Interpretando il Teorema di Cramer per i sistemi omogenei, si ottiene che se det$\mathbf{A} \neq 0$ allora $\mathbf{x} = \mathbf{0}$ è l'unica soluzione (infatti $\mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$).
Se invece det$\mathbf{A} = \mathbf{0}$, allora per il teorema sul rango si ha che le colonne di $\mathbf{A}$, indicate per esempio con $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ sono vettori linearmente dipendenti. Pertanto, per definizione di lineare dipendenza, esistono $x_1, x_2, \dots, x_n \in \mathbb{R}$ coefficienti non tutti nulli tali che

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}.$$

Posto $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ si ottiene

$$\mathbf{A}\mathbf{x} = \mathbf{0} \tag{1.3.1}$$

con $\mathbf{x} \neq \mathbf{0}$ perché i coefficienti x_i sono non tutti nulli. Quindi il sistema omogeneo (1.3.1) ha almeno una soluzione non nulla. D'altra parte, osservando che

$$\mathbf{A}(\lambda\mathbf{x}) = \lambda\mathbf{A}(\mathbf{x}) = \mathbf{0}$$

si vede che in realtà le soluzioni sono infinite. Si può allora enunciare il seguente risultato.

Teorema 1.3.1. *Il sistema lineare omogeneo di n equazioni in n incognite $\mathbf{A}\mathbf{x} = \mathbf{0}$ ha solo la soluzione banale $\mathbf{x} = \mathbf{0}$ se e soltanto se det$\mathbf{A} \neq 0$; ha almeno una soluzione non banale (in questo caso ne ha infinite) se e soltanto se det$\mathbf{A} = 0$.*

☞ **Osservazione 1.3.2.** Vediamo cosa succede al caso del sistema lineare non omogeneo se det$\mathbf{A} = 0$. Il caso generale verrà trattato più avanti col Teorema di Rouché-Capelli. Per il momento ci limitiamo ad osservare che se $\mathbf{x}_0$ è soluzione del sistema omogeneo $\mathbf{A}\mathbf{x}_0 = \mathbf{0}$ e $\mathbf{x}_1$ è soluzione del sistema non omogeneo $\mathbf{A}\mathbf{x}_1 = \mathbf{b}$ allora $\mathbf{x}_0 + \mathbf{x}_1$ è soluzione del sistema non omogeneo: infatti

$$\mathbf{A}(\mathbf{x}_1 + \mathbf{x}_0) = \mathbf{A}\mathbf{x}_1 + \mathbf{A}\mathbf{x}_0 = \mathbf{b} + \mathbf{0} = \mathbf{b}. \tag{1.3.2}$$

Quindi, se det$\mathbf{A} = 0$ e se $\mathbf{A}\mathbf{x} = \mathbf{b}$ ha una soluzione, ne ha infinite. Quindi se det$\mathbf{A} = 0$, a differenza del Teorema di Cramer, per il sistema lineare non omogeneo $\mathbf{A}\mathbf{x} = \mathbf{b}$ viene necessariamente a cadere l'esistenza o l'unicità: il sistema non omogeneo è impossibile o indeterminato.

✎ **Esempio 1.3.3.** *Sia $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^3$ la trasformazione lineare rappresentata nella base canonica da*

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

Determinare Ker$\mathcal{L}$ *e* Im$\mathcal{L}$, *le loro dimensioni e una loro base.*

Prima di tutto andiamo a calcolare il rango di $\mathbf{A}$ che ci dà la dimensione dell'immagine di $\mathcal{L}$. Si vede subito che il suo determinante è nullo, mentre isolando la matrice

$$\begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix}$$

esso ha determinante diverso da zero quindi il rango di $\mathbf{A}$ vale 2; quindi la dimensione di Im$\mathcal{L} = 2$ e, dal teorema di nullità più rango, la dimensione di Ker$\mathcal{L} = 3 - 2 = 1$.

Per determinare il nucleo di $\mathcal{L}$ occorre risolvere il sistema omogeneo

$$\begin{cases} 2x_1 + x_3 = 0 \\ 3x_1 + x_2 = 0 \\ -x_1 - x_2 + x_3 = 0. \end{cases}$$

La matrice associata ha rango 2, quindi basta considerare solo due delle precedenti equazioni (per esempio eliminando l'ultima equazione) e portare a destra uno dei tre valori di x_i, per esempio x_1 che sarà il parametro in funzione del quale verranno esplicitate le altre due componenti del generico vettore del nucleo. Quindi il precedente sistema risulta equivalente a

$$\begin{cases} x_3 = -2x_1 \\ x_2 = -3x_1 \end{cases}$$

quindi il generico vettore del nucleo è $(x_1, -3x_1, -2x_1)$ quindi una base del Ker$\mathcal{L}$ è

$$\begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}.$$

Per determinare l'immagine di $\mathcal{L}$ andiamo a cercare i vettori (y_1, y_2, y_3) per i quali il seguente sistema non omogeneo ammette soluzione (i vettori per i quali esiste una controimmagine nello spazio di partenza)

$$\begin{cases} 2x_1 + x_3 = y_1 \\ 3x_1 + x_2 = y_2 \\ -x_1 - x_2 + x_3 = y_3 \end{cases}$$

da cui si ricava esplicitamente che $y_1 = y_2 + y_3$. Quindi l'immagine di $\mathcal{L}$ è rappresentata dal piano di equazione $y_1 = y_2 + y_3$; il generico elemento dell'immagine pertanto si può scrivere come

$$\begin{pmatrix} y_2 + y_3 \\ y_2 \\ y_3 \end{pmatrix} = y_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

e quindi una base di tale sottospazio è data per esempio da

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Alternativamente si osserva che

$$\mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} x_3$$

e quindi una base dell'immagine è costituita da due vettori linearmente indipendenti, per esempio

$$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Si può facilmente verificare che questi due vettori generano lo stesso spazio vettoriale degli altri due scelti prima.

1.4. Sistemi lineari: il caso generale

Consideriamo ora il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.4.1}$$

di m equazioni e n incognite, e sia $\mathbf{A}$ la matrice dei coefficienti di tipo (m, n) che rappresenta una trasformazione lineare $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$, fissate delle basi in $\mathbb{R}^n$ e $\mathbb{R}^m$ rispettivamente. Il sistema lineare (1.4.1) si può riscrivere nel modo seguente

$$\mathbf{Ax} = \mathbf{b} \qquad \text{o anche} \qquad \mathcal{L}(\mathbf{x}) = \mathbf{b}.$$

Risolvere il sistema significa trovare quali $\mathbf{x} \in \mathbb{R}^n$ hanno come immagine tramite $\mathcal{L}$ il vettore $\mathbf{b} \in \mathbb{R}^m$ (detto altrimenti significa trovare le controimmagini di $\mathbf{b}$). Vale il seguente risultato generale.

Teorema 1.4.1. (TEOREMA DI ROUCHÉ-CAPELLI) *Sia* $\mathbf{A}$ *la matrice dei coefficienti del sistema* (1.4.1) *e sia* $\mathbf{B} = (\mathbf{A}|\mathbf{b})$ *la matrice completa ottenuta orlando la matrice* $\mathbf{A}$ *con la colonna dei termini noti. Allora il sistema ha soluzione se e soltanto se le matrici* $\mathbf{A}$ *e* $\mathbf{B}$ *hanno lo stesso rango.*

Il sistema per definizione ha soluzione se e soltanto se $\mathbf{b} \in \mathrm{Im}\mathcal{L}$. Quindi indicando con $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ le colonne di $\mathbf{A}$ e con $\mathbf{e}_j$ i vettori della base canonica di $\mathbb{R}^n$, si ha (teorema di rappresentazione) $\mathbf{A}\mathbf{e}_j = \mathbf{a}_j$, cioè le colonne di $\mathbf{A}$ sono i trasformati dei vettori della base canonica di $\mathbb{R}^n$. Di conseguenza, ogni elemento di $\mathrm{Im}\mathcal{L}$ è combinazione lineare delle colonne di $\mathbf{A}$. Il sistema per definizione è risolubile se e soltanto se $\mathbf{b} \in \mathrm{Im}\mathcal{L}$ se e soltanto se $\mathbf{b}$ è combinazione linare delle colonne di $\mathbf{A}$ se e soltanto se il rango della matrice $(\mathbf{A}|\mathbf{b})$ non è massimo se e soltanto se il rango di $(\mathbf{A}|\mathbf{b})$ coincide col rango di $\mathbf{A}$.

1.4.1. Il caso $n = m$

Interpretiamo il caso $n = m$. Se $\det\mathbf{A} \neq 0$ allora per forza le matrici $\mathbf{A}$ e $\mathbf{B}$ hanno lo stesso rango e pertanto il sistema è risolubile (dal Teorema di Cramer sappiamo anche che c'è unicità della soluzione). Se invece $\det\mathbf{A} = 0$, allora il rango di $\mathbf{A}$ è senz'altro strettamente minore di n ma il rango di $\mathbf{B}$ potrebbe essere eventualmente uguale a n: pertanto il sistema è impossibile o indeterminato.

1.4.2. Il sistema omogeneo

Se $\mathbf{b} = \mathbf{0}$ in (1.4.1), allora si ha sempre che il rango di $\mathbf{A}$ coincide con il rango di $\mathbf{B}$ pertanto il sistema è sempre risolubile. L'insieme delle soluzioni del sistema coincide con il $\mathrm{Ker}\mathcal{L}$ dove $\mathcal{L}$ è la trasformazione lineare rappresentata dalla matrice dei coefficienti $\mathbf{A}$. Quindi, indicato con r il rango di $\mathbf{A}$, dal teorema di nullità più rango si ha che la dimensione del $\mathrm{Ker}\mathcal{L} = n-r$. Quindi *le soluzioni del sistema omogeneo formano uno spazio vettoriale di dimensione* $n - r$. Si dice anche che il sistema ha ∞^{n-r} soluzioni per indicare che le soluzioni sono infinite e dipendenti da $n - r$ parametri arbitrari.

1.4.3. Il sistema non omogeneo

Tornando al sistema non omogeneo, ci chiediamo quante soluzioni ammette nel caso sia risolubile (cioè nel caso le matrici $\mathbf{A}$ e $\mathbf{B}$ abbiano lo stesso rango, come recita il Teorema di Rouché-Capelli. Come abbiamo visto dal conto (1.3.2), vale il seguente teorema.

Teorema 1.4.2. *Se il sistema non omogeneo* (1.4.1) *ammette una soluzione* $\mathbf{x}_1$ *allora le soluzioni del sistema non omogeneo sono tutte e sole le* n*-uple del tipo* $\mathbf{x}_1 + \mathbf{x}_0$ *al variare di* $\mathbf{x}_0 \in \mathrm{Ker}\mathcal{L}$ *(ossia al variare di* $\mathbf{x}_0$ *tra le soluzioni del sistema omogeneo).*

Pertanto anche il sistema non omogeneo ammette ∞^{n-r} soluzioni dove r è il rango della matrice dei coefficienti $\mathbf{A}$.

1.4.4. Cenni al metodo di riduzione

Per maggiori dettagli o approfondimenti si veda [L], Capitolo 2, paragrafo 2.3

1.5. Cambiamento di base e matrice rappresentativa

Consideriamo una trasformazione lineare $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^n$. La rappresentazione di $\mathcal{L}$ per mezzo della matrice $\mathbf{A}$ (in questo caso quadrata di tipo (n, n)) dipende dalla scelta della base di $\mathbb{R}^n$. È pertanto comprensibile il fatto che se $\mathbf{A}$ fosse "speciale" (per esempio diagonale o ridotta a scala) i calcoli (per la determinazione del rango ad esempio) sarebbero decisamente semplificati. Ci chiediamo dunque prima di tutto come cambia la rappresentazione di $\mathcal{L}$ cambiando la base di $\mathbb{R}^n$.
Supponiamo di essere in $\mathbb{R}^n$ con la base canonica $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (lo stesso procedimento si può applicare se $\mathbb{R}^n$ viene sostituito da un generico V spazio vettoriale di dimensione n). Notiamo che possiamo scegliere per comodità la stessa base per lo spazio di partenza e lo spazio di arrivo, visto che hanno la stessa dimensione.
Supponiamo ora di scegliere altri n vettori linearmente indipendenti, per esempio $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2 \ldots, \tilde{\mathbf{e}}_n$ e sia

$$\mathbf{S} = (\tilde{\mathbf{e}}_1 | \tilde{\mathbf{e}}_2 | \ldots | \tilde{\mathbf{e}}_n)$$

la matrice che ha come colonne i nuovi vettori. Si vede facilmente che

$$\mathbf{S}\mathbf{e}_i = \tilde{\mathbf{e}}_i \qquad i = 1, 2, \ldots, n.$$

Quindi per "definizione" di matrice rappresentativa, $\mathbf{S}$ ha per colonne le immagini dei vettori della base canonica tramite l'applicazione che fa passare da una base all'altra. Dunque $\mathbf{S}$ è la matrice che rappresenta, nella base canonica, la trasformazione che fa passare da una base all'altra (attenzione! non stiamo ancora rappresentando $\mathcal{L}$ nella nuova base!)
Visto che i nuovi vettori $\tilde{\mathbf{e}}_i$ sono linearmente indipendenti, allora $\det\mathbf{S} \neq 0$ quindi $\mathbf{S}$ è non singolare e pertanto invertibile. Allora

$$\mathbf{e}_i = \mathbf{S}^{-1}\tilde{\mathbf{e}}_i. \tag{1.5.1}$$

A questo punto allora, se un vettore di $\mathbb{R}^n$ ha componenti $(x_1, x_2, \ldots, x_n)$ rispetto alla base canonica, nella nuova base avrà coordinate

$$\tilde{\mathbf{x}} = \mathbf{S}^{-1}\mathbf{x} \qquad \tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n). \tag{1.5.2}$$

Infatti se

$$\begin{aligned}\mathbf{x} &= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \overset{(1.5.1)}{=} x_1\mathbf{S}^{-1}\tilde{\mathbf{e}}_1 + x_2\mathbf{S}^{-1}\tilde{\mathbf{e}}_2 + \cdots + x_n\mathbf{S}^{-1}\tilde{\mathbf{e}}_n \\ &= \mathbf{S}^{-1}x_1\tilde{\mathbf{e}}_1 + \mathbf{S}^{-1}x_2\tilde{\mathbf{e}}_2 + \cdots + \mathbf{S}^{-1}x_n\tilde{\mathbf{e}}_n = \tilde{x}_1\tilde{\mathbf{e}}_1 + \tilde{x}_2\tilde{\mathbf{e}}_2 + \cdots + \tilde{x}_n\tilde{\mathbf{e}}_n\end{aligned}$$

da cui la (1.5.2).

✎ **Esempio 1.5.1.** *Consideriamo* $\mathbb{R}^2$ *con la base canonica* $\mathbf{e}_1, \mathbf{e}_2$ *e sia* $\tilde{\mathbf{e}}_1 = (1,1)$ *e* $\tilde{\mathbf{e}}_2 = (2,0)$ *una nuova base di* $\mathbb{R}^2$*. Sia* $\mathbf{v} = (2,1)$ *rispetto alla base canonica. Allora*

$$\mathbf{S} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \qquad \mathbf{S}^{-1} = \begin{pmatrix} 0 & 1 \\ 1/2 & -1/2 \end{pmatrix}$$

Quindi

$$\tilde{\mathbf{x}} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$$

Pertanto

$$\mathbf{v} = (2,1) = 2\mathbf{e}_1 + 1\mathbf{e}_2 = \tilde{\mathbf{e}}_1 + \frac{1}{2}\tilde{\mathbf{e}}_2.$$

Sia ora $\mathbf{A}$ la matrice che rappresenta $\mathcal{L}$ rispetto alla base canonica e sia

$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

Allora se $\mathbf{y}$ sono le coordinate di un vettore di $\mathbb{R}^n$ rispetto alla base canonica e $\tilde{\mathbf{y}}$ sono le coordinate dello stesso vettore rispetto alla nuova base si ha

$$\tilde{\mathbf{y}} = \mathbf{S}^{-1}\mathbf{y} = \mathbf{S}^{-1}\mathbf{A}\mathbf{x} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\tilde{\mathbf{x}}.$$

Quindi abbiamo trovato che la trasformazione $\mathcal{L}$ nella nuova base è rappresentata dalla matrice

$$\tilde{\mathbf{A}} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}. \tag{1.5.3}$$

❒ **Definizione 1.5.2.** Si dice che due matrici legate dalla relazione (1.5.3) con $\mathbf{S}$ invertibile sono EQUIVALENTI O SIMILI.
Si dice che una matrice quadrata $\mathbf{A}$ a elementi in $\mathbb{K}$ è DIAGONALIZZABILE (su $\mathbb{K}$) se

$$\mathbf{A} = \mathbf{S}\Lambda\mathbf{S}^{-1}$$

con $\Lambda, \mathbf{S}$ opportune matrici a elementi in $\mathbb{K}$, Λ diagonale e $\mathbf{S}$ non singolare. La matrice $\mathbf{S}$ si dice MATRICE DI PASSAGGIO.

Quindi se esiste una base di $\mathbb{K}^n$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$) in cui la trasformazione che rispetto alla base canonica si rappresenta con la matrice $\mathbf{A}$ ora si rappresenta invece con una matrice diagonale (quindi se la matrice $\mathbf{A}$ è diagonalizzabile), abbiamo risolto il nostro problema.
Ci chiediamo dunque: data $\mathbf{A}$ matrice quadrata, è diagonalizzabile? Se sì, come si determina la corrispondente matrice diagonale e la matrice di passaggio?
Per fare questo abbiamo bisogno di introdurre le nozioni di autovalore e autovettore.

1.6. Autovalori e autovettori

Ci chiediamo ora se esistono vettori non nulli che vengono mandati in vettori paralleli a se stessi, cioè

$$\exists \mathbf{v} : \mathbf{A}\mathbf{v} = \lambda \mathbf{v} \qquad \text{per qualche } \lambda \in \mathbb{K}? \tag{1.6.1}$$

Lo scopo di questa richiesta è la seguente: è come chiedere che la trasformazione associata ad $\mathbf{A}$ sia di tipo diagonale. Infatti se esistono vettori $\mathbf{v}_i$ corrispondenti a certi valori λ_i come da (1.6.1), per $i = 1, \ldots, n$, allora è facile vedere che

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i \Leftrightarrow \mathbf{A}\mathbf{S} = \mathbf{S}\Lambda$$

con Λ la matrice diagonale con elementi λ_i sulla diagonale principale e $\mathbf{S}$ la matrice che ha $\mathbf{v}_i$ come colonne. Dunque dalla (1.5.3) $\mathbf{A}$ è simile a Λ.
Nel caso della (1.6.1) dunque per linearità si ha

$$\mathbf{A}(t\mathbf{v}) = t\mathbf{A}(\mathbf{v}) = (t\lambda)\mathbf{v} \tag{1.6.2}$$

pertanto se esiste $\mathbf{v}$ associato a λ in (1.6.1), allora ne esistono infiniti.
Scrivendo $\mathbf{I}_n\mathbf{v}$ al posto di $\mathbf{v}$ possiamo riscrivere l'equazione $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ come

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{v} = \mathbf{0} \qquad \mathbf{v} \text{ incognito (come } \lambda).$$

Le componenti di $\mathbf{v}$ soddisfano un sistema lineare omogeneo di n equazioni e n incognite. Dalla teoria dei sistemi lineari, sappiamo che, perché esistano vettori (non nulli!) del sistema omogeneo, si deve avere determinante nullo della matrice dei coefficienti associata, cioè nel nostro caso

$$\det(\mathbf{A} - \lambda\mathbf{I}_n) = |\mathbf{A} - \lambda\mathbf{I}_n| = 0. \tag{1.6.3}$$

Si tratta di un'equazione algebrica in λ di grado n.

❒ **Definizione 1.6.1.** L'equazione (1.6.3) si dice EQUAZIONE CARATTERISTICA; il polinomio

$$D(\lambda) = |\mathbf{A} - \lambda\mathbf{I}_n|$$

si chiama POLINOMIO CARATTERISTICO.

Dal teorema fondamentale dell'algebra si sa che questo polinomio ammette sempre esattamente n soluzioni, in campo reale o complesso. Si dimostra che $D(\lambda)$ non cambia se si sostituisce alla matrice $\mathbf{A}$ un'altra matrice equivalente, nel senso della (1.5.3).

❒ **Definizione 1.6.2.** Chiamiamo AUTOVALORE di $\mathbf{A}$ (o della trasformazione lineare che la matrice $\mathbf{A}$ rappresenta) qualsiasi $\lambda \in \mathbb{C}$ che soddisfi l'equazione caratteristica (1.6.3). Chiameremo AUTOVETTORE (corrispondente all'autovalore λ) ogni vettore $\mathbf{v} \neq \mathbf{0}$ con $\mathbf{v} \in \mathbb{K}^n$ che risolve l'equazione $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

☞ **Osservazione 1.6.3.** Come abbiamo visto dalla (1.6.2), ad ogni autovalore corrispondono infiniti autovettori.

✎ **Esempio 1.6.4.** *Determinare autovalori e autovettori della seguente matrice*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

assieme a una base per ciascun autospazio.

Prima di tutto andiamo a determinare gli autovalori di $\mathbf{A}$. Si ha

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix}.$$

Da cui

$$\det(\mathbf{A} - \lambda\mathbf{I}) = -\lambda^3 + 1 + 1 + \lambda + \lambda + \lambda = -\lambda^3 + 3\lambda + 2 = -(\lambda+1)^2(\lambda-2).$$

Quindi gli autovalori del sistema sono $\lambda = -1$ con molteplicità 2 e $\lambda = 2$ con molteplicità 1. Andiamo ora a calcolare gli autospazi relativi agli autovalori della matrice di partenza. Si ha

$$\mathbf{A} + \mathbf{I} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

quindi l'autospazio relativo all'autovalore $\lambda = -1$ è $x + y + z = 0$; una base dell'autospazio è

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

D'altra parte

$$\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

quindi l'autospazio relativo all'autovalore $\lambda = 2$ è dato dal sistema

$$\begin{cases} x - 2y = -z \\ x + y = 2z \end{cases}$$

che risolto (ad esempio con la regola di Cramer o con calcoli diretti) dà $x = y = z$. Una base dell'autospazio è dunque data da

$$\mathbf{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

❒ **Definizione 1.6.5.** Si dice che un autovalore λ_0 ha molteplicità algebrica m se λ_0 è una radice del polinomio caratteristico $D(\lambda)$ di molteplicità m, cioè se $D(\lambda)$ è divisibile per $(\lambda - \lambda_0)^m$ ma non per $(\lambda - \lambda_0)^{m+1}$.

☞ **Osservazione 1.6.6.** Se $\lambda_1, \lambda_2, \ldots, \lambda_k$ per $k \leq n$ sono gli autovalori distinti di $\mathbf{A}$ (reali o complessi) di molteplicità rispettivamente $m_1, m_2, \ldots, m_k$ allora per il teorema fondamentale dell'algebra si deve per forza avere $\sum_{i=1}^{k} m_i = n$.

❒ **Definizione 1.6.7.** L'insieme degli autovettori corrispondenti all'autovalore λ_j (unito al vettor nullo) costituisce il nucleo della trasformazione lineare $\mathbf{A} - \lambda_j \mathbf{I}_n$. Questo è uno spazio vettoriale chiamato AUTOSPAZIO associato all'autovalore λ_j. La dimensione dell'autospazio relativo all'autovalore λ_j è detta MOLTEPLICITÀ GEOMETRICA dell'autovalore λ_j (e la indicheremo con d_j).

Vale il seguente risultato.

Teorema 1.6.8.

- *Autovettori relativi ad autovalori distinti sono linearmente indipendenti*
- *La molteplicità geometrica di un autovalore non supera mai la sua molteplicità algebrica, in simboli* $d_j \leq m_j$.

❒ **Definizione 1.6.9.** Un autovalore λ_j si dice REGOLARE se la sua molteplicità algebrica coincide con la sua molteplicità geometrica, in simboli $d_j = m_j$.

☞ **Osservazione 1.6.10.** Se un autovalore è semplice (cioè $m_j = 1$) allora necessariamente anche $d_j = 1$ (perché per definizione gli autovettori sono diversi dal vettor nullo). Quindi *un autovalore semplice è sempre regolare.* La distinzione tra autovalori regolari o no ha dunque senso per autovalori di molteplicità algebrica maggiore di 1.

Teorema 1.6.11. (condizioni di diagonalizzabilità) *La matrice* $\mathbf{A}$ *è diagonalizzabile su* $\mathbb{K}$ *se e soltanto se* $\mathbb{K}^n$ *possiede una base di autovettori di* $\mathbf{A}$. *In tal caso, detti* $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ *gli autovettori di tale base e* $\lambda_1, \lambda_2, \dots, \lambda_n$ *gli autovalori corrispondenti (*$\lambda_i \in \mathbb{K}$ *non necessariamente distinti) risulta*

$$\mathbf{A} = \mathbf{S}\Lambda\mathbf{S}^{-1}$$

dove

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \qquad \mathbf{S} = (\mathbf{h}_1|\mathbf{h}_2|\dots|\mathbf{h}_n)$$

Se $\mathbf{A}$ è diagonalizzabile, allora $\mathbf{A} = \mathbf{S}\Lambda\mathbf{S}^{-1}$ con $\mathbf{S}$ non singolare e Λ diagonale. Allora $\mathbf{AS} = \mathbf{S}\Lambda$. Chiamiamo $\lambda_1, \lambda_2, \dots, \lambda_n$ gli elementi della matrice diagonale Λ e $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ le colonne di $\mathbf{S}$. Quindi $\mathbf{AS} = \mathbf{S}\Lambda$ è equivalente a scrivere $\mathbf{Ah}_i = \lambda_i\mathbf{h}_i$ quindi gli $\mathbf{h}_i$ sono autovettori relativi agli autovalori λ_i corrispondenti, per $i = 1, \dots, n$. D'altra parte $\mathbf{S}$ è non singolare, quindi $\mathbf{h}_i$ sono linearmente indipendenti ed effettivamente sono una base di autovettori.
Supponiamo ora che esista una base di autovettori $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ e siano $\lambda_1, \lambda_2, \dots, \lambda_n$ i corrispondenti autovalori (non necessariamente distinti). Se definiamo

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \qquad \mathbf{S} = (\mathbf{h}_1|\mathbf{h}_2|\dots|\mathbf{h}_n)$$

allora le relazioni $\mathbf{Ah}_i = \lambda_i\mathbf{h}_i$ si riscrivono come $\mathbf{AS} = \mathbf{S}\Lambda$ cioè $\mathbf{A} = \mathbf{S}\Lambda\mathbf{S}^{-1}$. La matrice $\mathbf{S}$ è invertibile perché gli $\mathbf{h}_i$ sono una base di autovettori, quindi sono linearmente indipendenti.

✎ **Esempio 1.6.12.** *Torniamo all'esempio precedente e consideriamo la matrice*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

di cui abbiamo determinato autovalori e autospazi relativi. Una base di autovettori è costituita *da*

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad \mathbf{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Quindi ponendo

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

si vede (come è naturale) che $\mathbf{S}$ *è non singolare quindi è possibile determinare la sua inversa*

$$\mathbf{S}^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

A questo punto allora si ottiene

$$\Lambda = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Teorema 1.6.13. (CONDIZIONI DI DIAGONALIZZABILITÀ) *La matrice* $\mathbf{A}$ *a elementi in* $\mathbb{K}$ *è diagonalizzabile su* $\mathbb{K}$ *se e soltanto se i suoi autovettori sono tutti regolari. In particolare una matrice con tutti gli autovalori semplici è sempre diagonalizzabile.*

1.7. Matrici reali simmetriche

❐ **Definizione 1.7.1.** Una matrice reale $\mathbf{A}$ di tipo (n, n) si dice ORTOGONALE se l'inversa coincide con la trasposta, cioè

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}_n$$

✎ **Esempio 1.7.2.**

$$\mathbf{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Teorema 1.7.3. (proprietà delle matrici ortogonali)

- $\mathbf{A}$ *è una matrice ortogonale di tipo* $(n.n)$ *se e soltanto se le sue righe sono una base ortonormale di* $\mathbb{R}^n$ *(idem per le colonne)*
- *Il determinante di una matrice ortogonale vale* 1 *o* -1
- *Il prodotto di matrici ortogonali è ancora ortogonale*
- *La trasformazione lineare rappresentata da una matrice ortogonale conserva il modulo dei vettori e il prodotto scalare, cioè*

$$|\mathbf{Ax}| = |\mathbf{x}| \quad \forall\, \mathbf{x} \in \mathbb{R}^n \qquad \mathbf{Ax} \cdot \mathbf{Ay} = \mathbf{x} \cdot \mathbf{y}$$

Se $n = 2, 3$ la trasformazione lineare rappresentata da una matrice ortogonale è una rotazione (eventualmente accompagnata da una riflessione, vedi file di Esercizi).

Teorema 1.7.4. (diagonalizzazione per passaggio ortogonale) *Sia* $\mathbf{A}$ *matrice reale simmetrica. Allora* $\mathbf{A}$ *è diagonalizzabile con una matrice di passaggio ortogonale, cioè*

$$\mathbf{A} = \mathbf{M}\Lambda\mathbf{M}^T$$

con Λ *diagonale e* $\mathbf{M}$ *ortogonale. Questo significa in particolare che* $\mathbb{R}^n$ *possiede una base ortonormale di autovettori di* $\mathbf{A}$.

✎ **Esempio 1.7.5.** *Trovare la matrice di passaggio ortogonale che diagonalizza la seguente matrice simmetrica*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Da esercizi precedenti abbiamo visto che gli autovalori di $\mathbf{A}$ sono $\lambda = -1$ con molteplicità 2 e $\lambda = 2$ con molteplicità 1; l'autospazio relativo all'autovalore $\lambda = -1$ è $x + y + z = 0$; una base dell'autospazio è

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Inoltre l'autospazio relativo all'autovalore $\lambda = 2$ è $x = y = z$ e una sua base è data da

$$\mathbf{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

A questo punto, per ottenere la matrice di passaggio $\mathbf{M}$ che diagonalizza la matrice di partenza, è necessario determinare tre autovettori ortonormali. Per fare questo è sufficiente determinare basi ortonormali dei due autospazi.
Relativamente all'autovalore $\lambda = 2$ è facile, basta normalizzare $\mathbf{w}_3$ e dunque un primo versore della base ortonormale cercata risulta essere

$$\mathbf{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Per quanto riguarda l'autospazio relativo all'autovalore $\lambda = -1$ intanto scegliamo uno dei due vettori della base data da $\mathbf{v}_1, \mathbf{v}_2$, per esempio scegliamo $\mathbf{w}_1$ e lo normalizziamo, dunque il secondo versore della base ortonormale cercata risulta essere

$$\mathbf{v}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

A questo punto, per determinare il terzo versore della base ortonormale cercata $\mathbf{v}_3$ basta cercare un vettore che sia combinazione lineare di $\mathbf{w}_1$ e $\mathbf{w}_2$ e ortogonale a $\mathbf{v}_2$, e poi normalizzarlo. Si deve dunque imporre $(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) \cdot \mathbf{v}_2 = 0$ (oppure si può procedere in maniera equivalente col metodo di ortonormalizzazione di Gram-Schmidt), cioè

$$\left(c_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right) \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 0 \quad \Leftrightarrow \quad c_1\sqrt{2} + \frac{c_2}{\sqrt{2}} = 0 \quad \Leftrightarrow \quad c_1 = -\frac{c_2}{2}.$$

Per esempio andiamo a scegliere $c_1 = 1$ e $c_2 = -2$. La base ortonormale cercata è dunque

$$\mathbf{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}.$$

La matrice di passaggio risulta essere

$$\mathbf{M} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

ed è una matrice non speciale (con determinante uguale a -1). A questo punto $\mathbf{M}^{-1} = \mathbf{M}^T$ e dunque $\mathbf{M}^{-1}\mathbf{AM}$ dà la matrice diagonale

$$\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

1.8. Esercizi di ricapitolazione

✎ **Esempio 1.8.1.** *Sia $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare rappresentata dalla matrice (rispetto alla base canonica)*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1+k^2 & 1+k \\ 2 & 4 & 3+k^2 \end{pmatrix}$$

1) Posto $k = 1$ calcolare autovalori di $\mathbf{A}$ e una base per ogni suo autospazio. $\mathbf{A}$ è diagonalizzabile?
2) Calcolare i valori di $k \in \mathbb{R}$ per cui $\mathcal{L}$ è suriettiva. Per tali valori $\mathcal{L}$ è iniettiva?

1) Se $k = 1$ allora

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & 2 \\ 2 & 4 & 4 \end{pmatrix}$$

e

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 1-\lambda & 2 & 2 \\ 1 & 2-\lambda & 2 \\ 2 & 4 & 4-\lambda \end{pmatrix}$$

quindi

$$\det(\mathbf{A}-\lambda\mathbf{I}) = (1-\lambda)(2-\lambda)(4-\lambda)+8+8-4(2-\lambda)-2(4-\lambda)-8(1-\lambda) = -\lambda^3+7\lambda^2 = 0 \Leftrightarrow \lambda^2(\lambda-7) = 0$$

quindi gli autovalori di $\mathbf{A}$ sono $\lambda = 0$ con molteplicità algebrica 2 e $\lambda = 7$ con molteplicità algebrica 1.
Calcoliamo i relativi autospazi. Se $\lambda = 0$, allora $\mathbf{A} - \lambda\mathbf{I} = \mathbf{A}$. Il rango di $\mathbf{A}$ è 1 quindi la dimensione dell'autospazio relativo all'autovalore $\lambda = 0$ è 2. Pertanto gli autovalori sono tutti regolari e la matrice $\mathbf{A}$ è diagonalizzabile.
L'autospazio relativo all'autovalore $\lambda = 0$ è il piano $x + 2y + 2z = 0$ e una sua base è

$$\mathbf{w}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

Consideriamo ora l'autovalore $\lambda = 7$. Si ha

$$\mathbf{A} - 7\mathbf{I} = \begin{pmatrix} -6 & 2 & 2 \\ 1 & -5 & 2 \\ 2 & 4 & -3 \end{pmatrix}$$

che ha rango 2 quindi la dimensione dell'autospazio relativo all'autovalore $\lambda = 7$ (come ci aspettavamo) è 1. Risolvendo ad esempio col metodo di Cramer (provare anche col metodo di riduzione!) si ottiene che l'autospazio cercato è la retta

$$\begin{pmatrix} \frac{z}{2} \\ \frac{z}{2} \\ z \end{pmatrix}$$

quindi una base ad esempio è

$$\mathbf{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Si ha che $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ costituiscono una base di autovettori (a conferma del fatto che $\mathbf{A}$ è diagonalizzabile).

2) Affinché $\mathcal{L}$ sia suriettiva, si deve avere che la dimensione dell'immagine di $\mathcal{L}$ è 3, quindi dal teorema di nullità più rango si ha che la dimensione del nucleo è 0, quindi $\mathcal{L}$ è suriettiva se e soltanto se $\mathcal{L}$ è iniettiva. I valori che rendono $\mathcal{L}$ suriettiva (e pertanto iniettiva) sono quelli diversi dai valori che annullano il determinante di $\mathbf{A}$, quindi

$$\begin{aligned} \det\mathbf{A} \neq 0 \quad &\Leftrightarrow \quad (1+k^2)(3+k^2) + 4(1+k) + 8 - 4(1+k^2) - 4(1+k) - 2(3+k^2) \neq 0 \\ &\Leftrightarrow \quad k^4 - 2k^2 + 1 \neq 0 (k^2-1)^2 \neq 0 \Leftrightarrow k \neq \pm 1 \end{aligned}$$

✎ **Esempio 1.8.2.** *Si consideri la matrice*

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & k \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

e l'applicazione lineare $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^3$ *definita da* $\mathcal{L}([x, y, z]) = \mathbf{A}(x, y, z)^T$.
a) Determinare al variare di $k \in \mathbb{R}$ *nucleo e immagine di* $\mathcal{L}$ *precisandone una base*
b) Posto $k = -1$ *stabilire se esiste una base in cui* $\mathcal{L}$ *è associata a una matrice diagonale*
c) Posto $k = -1$, *determinare se esistono le controimmagini di* $(2, -2, -1)$.

a) Lo spazio Im$\mathcal{L}$ è generato dalle colonne di $\mathbf{A}$. Le prime due sono uguali, la terza è indipendente per ogni valore di k, quindi la dimensione di Im$\mathcal{L}$ è 2 e una sua base è data da

$$\mathbf{w}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} k \\ 1 \\ 1 \end{pmatrix}$$

Il Ker$\mathcal{L}$ è dato dalle soluzioni del sistema lineare omogeneo

$$\begin{cases} kz = 0 \\ z = 0 \\ x + y + z = 0 \end{cases}$$

che danno

$$\begin{pmatrix} x \\ -x \\ 0 \end{pmatrix} \qquad x \in \mathbb{R}$$

quindi una sua base è data da

$$\mathbf{w}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

b) Se $k = -1$ allora

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

quindi

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -\lambda & 0 & -1 \\ 0 & -\lambda & 1 \\ 1 & 1 & 1-\lambda \end{pmatrix}.$$

Allora
$$\det(\mathbf{A}-\lambda\mathbf{I}) = \lambda^2(1-\lambda) - \lambda + \lambda = 0 \Leftrightarrow \lambda^2(1-\lambda) = 0$$
quindi gli autovalori di $\mathbf{A}$ sono $\lambda = 0$ con molteplicità algebrica 2 e $\lambda = 1$ con molteplicità algebrica 1.
D'altra parte, se $\lambda = 0$ allora
$$\mathbf{A}-\lambda\mathbf{I} = \mathbf{A} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$
che ha rango 2, quindi la dimensione dell'autospazio relativo è 1, pertanto la molteplicità geometrica dell'autovalore $\lambda = 0$ è $1 \neq 2$ che era la molteplicità algebrica. L'autovalore $\lambda = 0$ pertanto non è regolare e dunque la matrice $\mathbf{A}$ è diagonalizzabile.

c) Determinare le controimmagini del vettore $\mathbf{b}$ significa trovare gli $\mathbf{x}$ tali che $\mathbf{Ax} = \mathbf{b}$, quindi risolvere il sistema lineare non omogeneo $\mathbf{Ax} = \mathbf{b}$. Dal Teorema di Rouché-Capelli, il sistema è risolubile se e soltanto se il rango di $\mathbf{A}$ coincide con il rango della matrice completa (in questo caso, entrambi sono uguali a 2). Quindi le soluzioni richieste sono le soluzioni del sistema lineare omogeneo
$$\begin{cases} -z = 2 \\ z = -2 \\ x + y + z = -1 \end{cases}$$
che dà come soluzioni
$$\begin{pmatrix} x \\ 1-x \\ -2 \end{pmatrix} \qquad x \in \mathbb{R}$$
N.B.1: Si noti che le controimmagini di un dato vettore non necessariamente sono un sottospazio! (ad esempio nel nostro caso la retta trovata non passa per l'origine, quindi non è un sottospazio)
N.B.2: Determinare le controimmagini del vettore $\mathbf{b}$ significa trovare gli $\mathbf{x}$ tali che $\mathbf{Ax} = \mathbf{b}$, quindi risolvere il sistema lineare non omogeneo $\mathbf{Ax} = \mathbf{b}$. In linea di principio si può anche direttamente avere $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ ma **solo se la matrice A è non singolare**. Nel nostro caso la matrice $\mathbf{A}$ ha determinante nullo, quindi non può essere invertita.

✎ **Esempio 1.8.3.** *Sia $T : \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare data da*
$$T(1,1) = (1,2) \qquad T(1,2) = (1,1).$$
Determinare la matrice rappresentativa rispetto alla base canonica e stabilire se esiste una base rispetto alla quale la matrice rappresentativa è diagonale. In tal caso determinarla.

Si ha

$$(1,1) = \mathbf{e}_1 + \mathbf{e}_2 \qquad (1,2) = \mathbf{e}_1 + 2\mathbf{e}_2$$

quindi dalla linearità di T si ha

$$T(\mathbf{e}_1 + \mathbf{e}_2) = T(\mathbf{e}_1) + T(\mathbf{e}_2) = (1,2) \qquad T(\mathbf{e}_1 + 2\mathbf{e}_2) = T(\mathbf{e}_1) + 2T(\mathbf{e}_2) = (1,1)$$

quindi ponendo $T(\mathbf{e}_1) = (a,b)$ e $T(\mathbf{e}_2) = (c,d)$ e risolvendo il sistema corrispondente, si ottiene $T(\mathbf{e}_1) = (1,3)$ e $T(\mathbf{e}_2) = (0,-1)$. Pertanto la matrice rappresentativa rispetto alla base canonica è

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix}.$$

Alternativamente si poteva procedere nel seguente modo. Sia $\tilde{\mathbf{e}}_1 = (1,1)$ e $\tilde{\mathbf{e}}_2 = (1,2)$ una nuova base per $\mathbb{R}^2$ e sia

$$\mathbf{S} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad \text{da cui} \qquad \mathbf{S}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

A questo punto ricaviamo la matrice $\tilde{\mathbf{A}}$ rappresentativa rispetto alla nuova base: per definizione, questa matrice ha per colonne le immagini dei vettori $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2$ ma **rispetto alla nuova base!!!**. Cioè

$$T(\tilde{\mathbf{e}}_1) = \tilde{\mathbf{e}}_2 \qquad T(\tilde{\mathbf{e}}_2) = \tilde{\mathbf{e}}_1$$

quindi la matrice rappresentativa $\tilde{\mathbf{A}}$ rispetto alla nuova base è

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

quindi la matrice rappresentativa rispetto alla base canonica diventa

$$\mathbf{A} = \mathbf{S}\tilde{\mathbf{A}}\mathbf{S}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix}.$$

Vediamo ora se esiste una base rispetto alla quale la matrice rappresentativa è diagonale. Questo è equivalente a cercare di capire se $\mathbf{A}$ è diagonalizzabile. Determiniamo il polinomio caratteristico. Si ha

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 1-\lambda & 0 \\ 3 & -1-\lambda \end{pmatrix}$$

quindi

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (1-\lambda)(-1-\lambda) = 0$$

quindi gli autovalori sono $\lambda = \pm 1$ entrambi con molteplicità algebrica 1. Entrambi gli autovalori dunque sono regolari quindi la matrice $\mathbf{A}$ è diagonalizzabile. Determiniamo una base di autovettori.

Se $\lambda = 1$ allora

$$\mathbf{A} - \mathbf{I} = \begin{pmatrix} 0 & 0 \\ 3 & -2 \end{pmatrix}$$

che ha (come ci si aspetta) rango 1. L'autospazio risulta $3x - 2y = 0$ quindi una base risulta ad esempio

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Se $\lambda = -1$ allora

$$\mathbf{A} + \mathbf{I} = \begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix}$$

che ha di nuovo rango 1. L'autospazio risulta $x = 0$ quindi una base risulta ad esempio

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

La matrice di passaggio risulta

$$\mathbf{S} = \begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix} \qquad \text{da cui} \qquad \mathbf{S}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ -3 & 2 \end{pmatrix}$$

e

$$\Lambda = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

✎ **Esempio 1.8.4.** *Risolvere il seguente sistema lineare al variare del parametro $k \in \mathbb{R}$*

$$\begin{cases} x + y + kz = 2 \\ x + y + 3z = k - 1 \\ 2x + ky - z = 1 \end{cases}$$

La matrice dei coefficienti è

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & k \\ 1 & 1 & 3 \\ 2 & k & -1 \end{pmatrix}.$$

mentre la colonna dei termini noti è

$$\begin{pmatrix} 2 \\ k - 1 \\ 1 \end{pmatrix}.$$

Si ha

$$\det\mathbf{A} = k^2 - 5k + 6.$$

Distinguiamo i vari casi.
$\boxed{k \neq 2, 3}$
Allora il determinante di $\mathbf{A}$ non si annulla e dal Teorema di Cramer esiste un'unica soluzione. Si ha

$$x = \frac{\begin{vmatrix} 2 & 1 & k \\ k-1 & 1 & 3 \\ 1 & k & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & k \\ 1 & 1 & 3 \\ 2 & k & -1 \end{vmatrix}} = \frac{k(k+2)}{k-2} \quad y = \frac{\begin{vmatrix} 1 & 2 & k \\ 1 & k-1 & 3 \\ 2 & 1 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & k \\ 1 & 1 & 3 \\ 2 & k & -1 \end{vmatrix}} = \frac{2(k+2)}{k-2} \quad z = \frac{\begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & k-1 \\ 2 & k & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & k \\ 1 & 1 & 3 \\ 2 & k & -1 \end{vmatrix}} = -1$$

$\boxed{k = 2}$
Allora il rango di $\mathbf{A}$ è 2, mentre il rango della matrice completa $\mathbf{B}$ è 3, quindi il sistema non è risolubile.
$\boxed{k = 3}$
Il rango della matrice $\mathbf{A}$ coincide con il rango della matrice $\mathbf{B}$ ed è 2, quindi il sistema ammette ∞^1 soluzioni, determinabili risolvendo il sistema

$$\begin{cases} x + y = 2 - 3z \\ 2x + 3y = 1 + z \end{cases}$$

Si ottiene

$$\begin{pmatrix} 5 - 10z \\ 7z - 3 \\ z \end{pmatrix} \qquad z \in \mathbb{R}.$$

✎ **Esempio 1.8.5.** *Sia data la matrice*

$$\mathbf{A} = \begin{pmatrix} k & 0 & 0 \\ 3 & 1 & 3 \\ k & 0 & 2 \end{pmatrix}.$$

Determinare per quali k è diagonalizzabile e determinare la matrice di passaggio.

Il polinomio caratteristico risulta

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (k - \lambda)(1 - \lambda)(2 - \lambda)$$

quindi gli autovalori sono $\lambda = k$, $\lambda = 1$ e $\lambda = 2$. A questo punto:
$\boxed{k \neq 1, 2}$

gli autovalori sono tutti distinti, per cui sono semplici e regolari, allora la matrice $\mathbf{A}$ è diagonalizzabile. La matrice di passaggio risulta in questo caso (le colonne sono una base di autovettori)

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & k-2 \\ 1 & 3 & 6 \\ 0 & 1 & k \end{pmatrix}$$

Si nota che essendo nel caso $k \neq 1, 2$ i tre vettori trovati sono indipendenti.
$\boxed{k=1}$
In tal caso gli autovalori sono $\lambda = 1$ con molteplicità algebrica 2 e $\lambda = 2$ con molteplicità algebrica 1. Se $\lambda = 1$ allora la dimensione dell'autospazio relativo è 2, pertanto tutti gli autovalori sono regolari e $\mathbf{A}$ è diagonalizzabile. L'autospazio relativo risulta $x + z = 0$, mentre l'autospazio relativo all'autovalore $\lambda = 2$ risulta $x = 0, y = 3z$. Quindi in questo caso la matrice di passaggio risulta

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ -1 & 0 & 1 \end{pmatrix}.$$

$\boxed{k=2}$
In tal caso gli autovalori sono $\lambda = 1$ con molteplicità algebrica 1 e $\lambda = 2$ con molteplicità algebrica 2. Se $\lambda = 2$ allora la dimensione dell'autospazio relativo è 1, quindi in questo caso la molteplicità algebrica dell'autovalore $\lambda = 2$ risulta diversa dalla sua molteplicità geometrica. L'autovalore non è regolare e la matrice $\mathbf{A}$ non è diagonalizzabile.

1.9. Complementi: elementi di geometria analitica nel piano e nello spazio

1.9.1. Equazione parametrica e cartesiana di una retta

Una retta nello spazio è individuata da:
- un punto e un vettore direzione;
- due punti;
- due piani non paralleli di cui è intersezione.

Per esempio in $\mathbb{R}^3$ (ma questo vale anche più in generale in $\mathbb{R}^n$) l'equazione parametrica vettoriale di una retta per $P_0 = (x_0, y_0, z_0)$ e di vettore direzione $\mathbf{v} = (a, b, c)$ è:

$$\begin{aligned} x &= x_0 + ta \\ y &= y_0 + tb \\ z &= z_0 + tc \qquad P = P_0 + t\mathbf{v} \end{aligned}$$

Se $a \neq 0$, $b \neq 0$, $c \neq 0$ allora eliminando il parametro t si ottiene l'equazione cartesiana della retta in $\mathbb{R}^3$

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Se uno dei parametri a, b, c è nullo, per esempio $a = 0$ si ha ad esempio l'equazione

$$x = x_0 \qquad \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

1.9.2. Equazione parametrica e cartesiana di un piano

Un piano in $\mathbb{R}^3$ può essere descritto ad esempio da un punto $P_0 = (x_0, y_0, z_0)$ e da un vettore direzione $\mathbf{n}$ ortogonale al piano. Per cui, se $P = (x, y, z)$ è il generico punto del piano, si ha che il vettore PP_0 è parallelo al piano, dunque (esprimendo questa condizione di parallelismo)

$$PP_0 \cdot \mathbf{n} = 0 \qquad \text{cioè} \qquad a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

e ponendo $d = ax_0 + by_0 + cz_0$ si ottiene

$$ax + by + cz = d.$$

Questa è l'equazione cartesiana del piano in $\mathbb{R}^3$. Per ricavare un'equazione parametrica ad esempio si può porre $x = s$, $y = t$ e ricavare z di conseguenza, in funzione di s e t. Ricordiamo che in $\mathbb{R}^3$ due equazioni cartesiane individuano una retta (a meno che non si tratti di due piani paralleli); un'equazione cartesiana individua un piano.

Per maggiori dettagli e approfondimenti, si veda [L], Capitolo 2, Paragrafo 2.5.